基于多子块联合估计的相关滤波跟踪 *

曹 洁[a, b]，解博江[a]，李 伟[b]，王进花[b]

(兰州理工大学 a. 计算机与通信学院; b. 电气工程与信息工程学院, 兰州 730050)

摘 要: 针对遮挡情况下相关滤波算法跟踪精度下降的问题，提出了一种基于多子块联合估计的核相关滤波跟踪方法。首先依据初始帧跟踪框的几何特征对目标自适应分块，并采用 KCF 方法对各子块独立跟踪得到联合置信图；然后以上帧目标的位置及尺度作为先验信息对搜索区域采样，同时将样本框中置信图的权值密度作为观测值，利用粒子滤波算法实现候选目标的最优估计；最后对置信度较低的子块反向投影至上帧图像进行遮挡检测，防止模板错误更新。定性和定量实验结果表明，该方法与原始 KCF 算法相比跟踪精度提升约 10%，具有良好的抗遮挡性，并对目标尺度变化具有一定的估计能力。

关键词: 分块跟踪；相关滤波；粒子滤波；遮挡检测；尺度变化

中图分类号: TP391.4 doi: 10.3969/j.issn.1001-3695.2017.07.0667

Correlation filter object tracking based joint multi-blocks estimation

Cao Jie[a, b], Xie Bojiang[a], Li Wei[b], Wang Jinhua[b]

(*a. College of Computer & Communication*, *b. College of Electrical & Information Engineering*, *Lanzhou University of Technology*, *Lanzhou* 730050, *China*)

Abstract: In order to solve the problem that the poor tracking accuracy of correlation filter algorithm under occlusion condition, this papre proposed a kernelized correlation filter tracking method using joint multiple blocks estimation. Firstly, it divided the target into several blocks adaptively according to the geometric features of the initial frame tracking box, and each block using the KCF method for tracking independently to get a combined confidence map. Then it took the location and scale of previous frame target as priori information sampling the search area, meanwhile, the weight density of the confidence map in the sample box is used as the observation value, achieve optimal estimation the candidate target using particle filter algorithm. Finally, blocks with lower confidence levels back-project to previous frame for occlusion detection to prevent template update mistakenly. The qualitative and quantitative experiment results shows that compared with the original KCF algorithm, the tracking accuracy of the proposed method improves about 10% and it is robust to occlusion and scale change in some degree.

Key Words: part-based object tracking; correlation filter; particle filter; occlusion detection; scale change

0 引言

视频目标跟踪是计算机视觉与模式识别的热点研究课题之一，在行为识别、智能监控、智能交通、汽车导航以及人机交互等领域扮演着十分重要的角色[1]。近年来，基于在线学习理论的判别式跟踪算法相比传统的生成式方法如 MeanShift[2]、粒子滤波[3]等拥有更高的跟踪精度与运行速度，引起了研究者的极大关注。

判别式方法将目标跟踪看成一个典型的分类问题，即如何将图像中的目标与背景进行区分，并且在对目标建模时充分利用了目标与背景信息，故具有较高的跟踪鲁棒性[4]。相关滤波是一种经典的判别式跟踪方法，Mosse[5]使用最小二乘训练滤波模板将相关滤波的概念首次引入目标跟踪领域，极大地提高了跟踪的实时性。STC[6]方法以概率论的方式在相关滤波算法中融入空时文本信息，并利用单个相关响应置信图实现了对目标尺度的有效估计。CSK[7]方法使用核化岭回归的方式训练滤波器模板，增强了模板在非线性情况下的有效性，通过构建循环矩阵巧妙地简化模板训练求解过程。KCF[8]方法在 CSK 的基础上，运用高斯核函数将单通道相关滤波器扩展为多通道，使其能够使用较为复杂的特征对目标进行更精确的建模，从而提高了跟踪精度。基于分块思想改进的相关滤波类跟踪方法，在目标被部分遮挡的情况下仍能利用未被遮挡部分提供的定位信息

基金项目: 国家自然科学基金资助项目（61263031）；甘肃省自然科学基金资助项目（1506RJZA105）

作者简介: 曹洁（1966-），女，安徽宿州人，教授，硕士，主要研究方向为信息融合、智能信息处理、模式识别；解博江（1991-），男，硕士研究生，主要研究方向为目标跟踪（603443495@qq.com）；李伟（1982-），男，博士，主要研究方向为目标跟踪、非线性滤波；王进花（1978-），副教授，博士，主要研究方向为故障诊断、智能信息处理.

进行跟踪，成为提升跟踪器抗遮挡能力的重要手段之一；而且相关滤波算法的高速处理能力为分块跟踪的实时性提供了可能[9]。文献[10]依据目标轮廓进行分块，采用 KCF 方法对子块独立跟踪，利用子块间定位结果的空间分布信息确定整体目标的位置及尺度。文献[11]在对子块进行独立的 CSK 跟踪的基础上，以子块跟踪置信图的主旁瓣比作为遮挡的判定依据对模板进行更新。上述基于多分块改进的跟踪方法仍有不足，首先未考虑子块可能存在漂移现象而直接依据分块跟踪结果所构成的空间几何特性对目标进行定位是不合理的，其次在子块模板更新中仅将置信图峰值的主旁瓣比（peak to sidelobe ratio, PSR）作为判据不能将外部遮挡与自身形变区分开来。

由此本文在对目标自适应分块及采用 KCF 方法对子块进行独立跟踪的基础上，为解决多子块线索融合中目标定位及尺度估计难问题，使用粒子滤波算法对多子块联合置信图进行采样，并依据采样框所包围形成的权值密度设定粒子权重，实现目标位置及尺度的最优估计；为了防止模板因遮挡而错误更新，将 PSR 较低子块的主分量直方图反向投影至上帧目标及背景区域，分别统计其像素投影概率，以此对遮挡情形进行判定。

1　分块策略和 KCF 跟踪原理

1.1　自适应分块策略

以固定形式对图像进行分块操作往往无法灵活应对跟踪目标的多样性，而初始帧目标跟踪框位置和大小一般由人为标定得来，在一定程度上能反映目标的整体轮廓。由此可以根据初始帧矩形跟踪框的几何特征对不同目标进行分块处理。当初始帧跟踪框横纵比大于设定阈值 T_1 时，将其在水平方向上切割为两个等面积的子块；当横纵比小于阈值 T_2 时，对跟踪框进行垂直分割；比值介于两阈值间的跟踪框将其等分为四个子块，如图 1 所示。

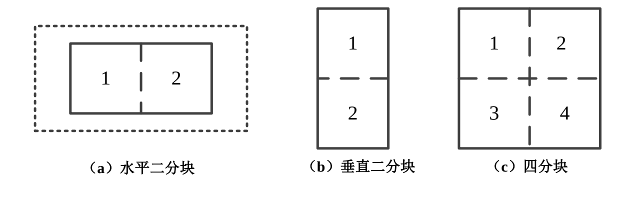


图 1　目标分块示意图

在对各子块进行跟踪时，为了充分利用背景信息同时减少潜在的相似目标干扰，本文将各子块的搜索区域限制为整体跟踪框面积大小的 2 倍，如图 1（a）中虚线框所示。

1.2　KCF 跟踪原理

KCF 跟踪方法对第 t-1 帧图像块 x 进行循环移位操作生成目标在不同位置处的样本集合 $\{x_1, x_2, ..., x_n\}$，改善在线学习跟踪过程中训练样本匮乏的问题。对目标的不同位移样本集，可以采用岭回归的形式在均方损失函数结构风险最小化的准则下训练求解出最优的滤波模板 w，使得线性分类模型 $y = w^T x$ 在所有的样本集合中均方误差和最小，即

$$\min_w \sum_{i=1} (w^T x_i - y_i)^2 + \lambda \|w\|^2 \tag{1}$$

在跟踪过程中，样本集合的特征空间往往是线性不可分的，KCF 使用高斯核函数 $\varphi(x)$ 将样本特征映射到高维空间中使其线性可分，增强了滤波模板在非线性情况下的适用性。在核空间中根据表现定理（representation theorem）[12],模板 w 可由所有的样本线性表示：

$$w = \sum_j a_j \varphi(x_j) \tag{2}$$

将其带入式（1）可得滤波模板的核化岭回归训练过程：

$$\min_a \sum_{i=1} (y_i - \sum_j a_j \varphi(x_j) \odot \varphi(x_i))^2 + \lambda \|\sum_j a_j \varphi(x_j)\|^2 \tag{3}$$

定义两个核函数点积 $\varphi(x) \odot \varphi(x') = \kappa(x, x')$ 并求解式（3）可以得到模板的闭合解：

$$a = (K + \lambda \mathbf{I})^{-1} \mathbf{Y} \tag{4}$$

其中：$\mathbf{Y}$ 是由元素 y_i 所组成的向量；I 为单位矩阵；K 则为由元素 $k_{ij} = \kappa(x_i, x_j)$ 所组成的核矩阵。如果直接通过式（4）求解系数 a，计算量较大。文献[7]已证明由循环样本集经核函数映射所组成的矩阵 K 仍为循环结构，充分利用任意循环矩阵都可以被傅里叶变换矩阵对角化这一特殊性质，如式（5）所示，简化了训练求解过程。

$$K = C(\mathrm{k}) = \mathbf{F} diag(k) \mathbf{F}^{\mathbf{H}} \tag{5}$$

其中：$\mathbf{F}$ 为傅里叶系数矩阵；k 是由元素 $k_i = \kappa(x, x_i)$ 所构成的向量；$diag(k)$ 为以向量 k 中的元素构成的对角阵；$C(k)$ 表示为以向量 k 构成的循环矩阵。式将（5）带入式（4），利用循环矩阵的性质可解得

$$a = C\left(f^{-1}\left(\frac{1}{\kappa(\hat{\mathbf{x}}, \hat{\mathbf{x}}') + \lambda}\right) \right) \odot y \tag{6}$$

对式两端同时进行 FFT 变换，并应用循环矩阵性质可得到所训练滤波模板的闭合解。

$$\hat{a} = \frac{\hat{\mathbf{y}}}{\kappa(\hat{\mathbf{x}}, \hat{\mathbf{x}}') + \lambda} \tag{7}$$

当第 t 帧图像到来后，以 t-1 帧图像中目标所在坐标为中心提取基本样本 z，同样以循环移位的方式并由快速傅里叶变换至频域计算每个样本的响应值，生成当前帧目标所在位置的置信图，其最大值处即为目标所在位置。响应过程如式（8）所示，其中 h 的元素为 $h_i = \kappa(x, z_i)$。

$$M^t = f^{-1}(\hat{a} \odot \hat{h}) \tag{8}$$

2　本文方法

首先根据初始帧跟踪框的几何特征对目标进行自适应分块；然后对每个子块采用 KCF 方法进行独立跟踪，生成各个子块目标位置的置信图及其权重。在此基础上加入粒子滤波算法对多子块联合置信图中目标的位置及尺度进行最优估计；融入遮挡检测模块以此实现子块模板的差异化更新。本文方法整体框架如图 2 所示。

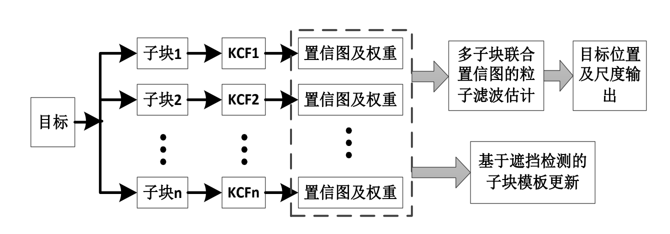


图 2　本文方法流程框架

2.1　多子块联合置信图的粒子滤波估计

以分块方式对目标进行跟踪时，往往认为子块跟踪结果所形成几何图案的质心即为目标所在中心点。这种判定方法未考虑子块可能由于遮挡等原因发生漂移，致使目标无法准确定位，而且不能针对目标的尺度变化进行估计。本文从 STC 算法的对单个相关响应置信图尺度估计过程中受到启发，拓展状态估计量，并采用粒子滤波算法对多子块联合置信图中目标大小及中心点坐标进行估计。

对目标子块采用 KCF 方法进行独立跟踪可得各子块置信图 M_j^t，其跟踪结果的可靠性根据文献[13]可由各子块置信图的 PSR 评判，则各子块归一化权重计算如式（9）所示。其中：G_j^t 为第 j 个子块置信图的峰值，μ_j^t 和 σ_j^t 为旁瓣的均值和标准偏差。

$$w_j^t = \frac{PSR_j^t}{\sum_{i=1} PSR_j^t}, PSR_j^t = \frac{G_j^t - \mu_j^t}{\sigma_j^t} \tag{9}$$

考虑到同一子块置信图中，权重应从最大峰值处到周边区域随距离逐级递减；并且由于目标可能存在尺度变化会造成多子块边界范围出现交错，其置信图也会发生层叠；而在多子块置信图加权融合过程中较大范围的边缘重叠会给主峰值带来干扰。为了缓解这种边缘效应，在多子块联合置信图计算过程中加入了一个余弦窗 C_j^t，该窗中心位置取决于相应子块置信图中最大峰值处，其范围由子块目标大小决定，最终使得中心处获得较高权重而边缘处像素权重趋于零。则多子块联合置信图为

$$M^t = \sum_{j=1}^{n} M_j^t w_j^t C_j^t \tag{10}$$

粒子滤波跟踪方法首先结合目标运动仿射信息建立状态转移方程，对粒子集的状态进行预测与约束；然后使用最新观测值构成的似然函数设定粒子集权重，将其由低似然区向高似然移动。在状态空间模型中构造如下所示的状态转移方程和似然函数：

$$x_t = Ax_{t-1} + V_x \tag{11}$$

$$p(y_t | x_t) = \frac{1}{\sqrt{2\pi\delta}} \exp(\frac{\rho_{x_t}}{2\delta^2}) \tag{12}$$

其中：$x_t = [x, y, w, h]^T$ 为 t 时刻目标的隐状态，对应着候选目标的跟踪框中心坐标以及尺度大小；A 为由仿射变换构成的转移矩阵；V_x 则为状态方程对应的高斯噪声协方差矩阵。ρ_{x_t} 计算方式如式（13）所示。其中：$M^t(x_t)$ 定义为 x_t 状态所对应的跟踪框在多子块联合置信图中包围权重的大小；s 是跟踪框所围像素的面积；ρ_{x_t} 为跟踪框中置信图权重的密度。

$$\rho_{x_t} = \frac{1}{s} M^t(x_t) \tag{13}$$

依据蒙特卡洛思想和大数定律，目标真实状态的后验概率 $p(x_{0t} | y_{1t})$ 可由对其采样的 N 个粒子组成的状态集联合归一化后的权重去逼近，即

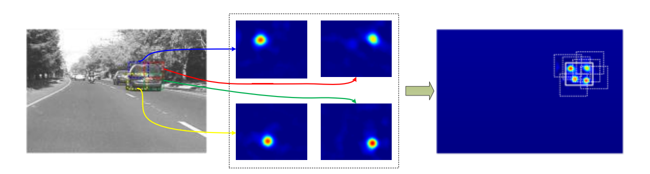
$$P(x_{0t} | y_{1t}) \approx \sum_{i=1}^{N} \omega_t^i \delta(x_{0t} - x_{0t}^i) \tag{14}$$

但实际应用中，直接从后验概率抽取粒子状态集较为困难，文献[14]使用次优建议分布函数 $p(x_t | x_{t-1})$ 代替真实后验概率进行重要性采样，其粒子集权重更新过程为

$$\omega_t^i \propto \omega_{t-1}^i p(y_k | x_t^i) \tag{15}$$

则根据最大后验准则 t 时刻目标状态的最优估计为

$$\hat{x}_t = \sum_{i=1}^{N} \omega_t^i x_t^i \tag{16}$$



(a)分块目标　　(b)多子块置信图　　(c)目标定位及尺度估计

图 3　分块目标跟踪流程

图 3 为分块目标定位及尺度估计过程。其中粒子滤波算法对多子块置信图的联合估计流程为：

1) 粒子初始化：设定粒子数 N，根据目标状态先验信息生成随机采样粒子集 $\{x_{0t}^i, \omega_1^i\}_{i=1}^N$，其相应的权值为 $\omega_1^i, i = 1, 2, \cdots, N$，并且满足 $\sum_{i=1}^{N} \omega_1^i = 1$

for　t=2:T　(T 为视频总帧数）

2）多子块置信图联合：根据各子块独立跟踪置信图 M_j^t 及其权重 w_j^t，结合窗函数 C_j^t 由（10）式生成联合置信图 M^t（其中余弦窗 C_j^t 中子块目标大小由 t-1 帧目标整体尺度估计结果与初始帧分块策略确定）；

3）状态预测：对一步状态转移方程（11）进行重要性采样，得到 t 时刻粒子集状态的预测值 x_t^i；

4）粒子权重更新：依据粒子集的预测状态，利用式（16）、（13）计算各粒子包围形成的置信图权值密度 $\rho_{x_t^i}$，并由式（15）得到 t 时刻粒子集所对应的权重 ω_t^i；

5）目标状态估计：对 ω_t^i 进行归一化处理并利用式（16）对目标的真实状态进行估计，得到候选目标位置及尺度的最优估计；

6）重采样：根据 ω_t^i 的大小，从粒子状态集 x_t^i 中随机选择 N 个粒子，生成重采样粒子集 $\{\tilde{x}_t^i, 1/N\}_{i=1}^N$；

end

2.2　基于遮挡检测的子块模板更新

遮挡及自身形变均会引起子块跟踪结果置信度的下降，而由外部遮挡引起的目标表观变化，遮挡物往往来自于背景所在区域，自身形变则来自于目标区域的可能性较高，因此本文在子块跟踪结果置信度较低的情况下，对该子块反向投影至上帧图像中的搜索区域，得到其像素值所属目标与背景区域的概率，进而检测目标是否被遮挡。

若在 t 帧，子块 q 置信度小于阈值，则该子块的归一化直方图如式（17）所示。其中：n 为子块 q 中的像素总数；$C_{q_t}(x,y)$ 表示 t 帧时子块 q 中 (x,y) 处像素的值（彩色目标为 HSV 空间中的色度分量，灰度目标则为像素灰度值）；δ 为克罗内克函数；u 是直方图 bin 值索引。

$$q(u)=\frac{1}{n}\sum_{(x,y)\in q}\delta[C_{q_t}(x,y)-u] \tag{17}$$

实际跟踪中，由于子块直方图分量较多，如果直接使用全部直方图进行反向投影，计算量较高，而且容易受到背景分量干扰；考虑到遮挡物及目标在搜索区域比重较大且像素值变化缓慢，所以采用直方图分量较大且空间分布较为集中的区域进行反向投影，选取直方图 $q(u)$ 中的 5 个最大分量组成 $\hat{u}$ 并分别计算其空间散度 σ_u 。

$$\begin{cases} m=\sum\limits_{(x,y)\in q}\delta(C_{q_t}(x,y)-\hat{u}) \\ \mu_u=\frac{1}{m}\sum\limits_{(x,y)\in q}\|(x,y)-(x_t,y_t)\|\,\delta(C_{q_t}(x,y)-\hat{u}) \\ V_u=\frac{1}{m}\sum\limits_{(x,y)\in q}(x_i-\mu_u)^T(x_i-\mu_u)\delta(C_{q_t}(x,y)-\hat{u}) \end{cases} \tag{18}$$

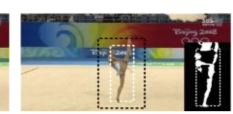
$$\sigma_u=\|V_u\| \tag{19}$$

其中：m 为子块 q 中所属 $\hat{u}$ 分量的像素总数，(x_t,y_t) 为当前帧目标所在位置的坐标。依据 σ_u 从 $\hat{u}$ 中选择空间分布较为集中即空间散度较小的 3 个分量构成 u'，并将其反向投影至上帧搜索区域得到像素概率分布图 $I_{t-1}(x,y)$ 。

$$I_{t-1}(x,y)=\sum_{(x,y)\in p}q(u')\delta[C_{q_{t-1}}(x,y)-u'] \tag{20}$$




(a)外部遮挡及其反向投影结果

(b)自身形变及其反向投影结果

图 4　不同情形的反向投影结果

当目标出现外部遮挡时，其反向投影的前景像素主要集中在背景区域的遮挡物上，如图 4（a）所示；而由于自身原因导致的表观变化，其投影结果主要集中在目标所在的区域，如图 4（b）所示。根据第 t-1 帧跟踪框位置，在反向投影图中确定目标及搜索区域，并分别统计目标区域前景像素个数 A_o 以及整个搜索区域前景像素个数 A_t，则子块 q 中像素反向投影至目标区域的概率为

$$P_o=\log\frac{A_o}{A_t-A_o} \tag{21}$$

若 P_o 小于遮挡判定阈值 P_{th}，说明子块中目标像素来自上一帧图像中的背景区域概率较大，即该子块目标出现背景遮挡，此时应停止其模板更新，相反则正常更新模板。则子块差异化模板更新过程为

$$\begin{cases} \hat{a}_{i,t}=(1-\varepsilon_{i,t}\eta)\hat{a}_{i,t-1}+\varepsilon_{i,t}\eta\hat{a}_{i,t} \\ \hat{X}_{i,t}=(1-\varepsilon_{i,t}\eta)\hat{X}_{i,t-1}+\varepsilon_{i,t}\eta\hat{X}_{i,t} \\ \varepsilon_{i,t}=\begin{cases}0 & P_{i,t}<P_{th} \\ 1 & P_{i,t}>P_{th}\end{cases} \end{cases} \tag{22}$$

3　实验与分析

3.1　实验参数与评价指标

实验部分所依托的硬件平台配置为：4.2 GHz 四核 Core i7-7700k 处理器，16 GB DDR3 内存；算法运行环境为 windows7+MATLAB R2014a。分块阈值 T_1 、T_2 是根据初始帧目标横纵比并通过遍历视频库所得到的经验值，实验中其值分别为 1.6 和 0.4；核化相关滤波器参数与 KCF 相同；目标状态估计过程中粒子数为 50，状态噪声服从高斯分布：$V_x\sim10^{-4}\times N(0.15,0.15,0.01,0.01)$；遮挡判定阈值 P_{th} 为 0.8。

为了对本文所提方法的有效性进行充分评估，采用文献[15]所建立的标准评价体系，以中心误差（center location error, CLE）、距离精度（distance precision, DP）以及重叠精度（overlap precision, OP）为指标，将本文方法与多种常用跟踪算法进行了定性和定量对比分析实验。CLE 是算法标记目标中心位置与真实值之间的欧氏距离；DP 则是 CLE 小于某一设定阈值（20 pixel）帧数占视频序列总长度的比值，能够反映跟踪过程中心点的定位精度；而 OP 是指覆盖域（overlap region,OR）得分大于设定阈值（0.5）帧数占视频序列总长度的比，在一定程度上能反映出跟踪框大小与目标的贴合程度，其中覆盖域得分计算为：$\text{score}=\frac{area(B_t\cap B_g)}{area(B_t\cup B_g)}$，$B_t$ 为第 t 帧时的跟踪框，B_g 为对应的目标真实值。

3.2　定性分析实验

算法定性分析实验选取了 Benchmark[15]测试集中的 2960 帧视频序列，包含遮挡、尺度放缩等目标跟踪中常见的挑战场景。将本文方法与 KCF、CSK 等近几年提出的经典相关滤波类算法在同一视频序列中进行对比分析。

遮挡场景分析:图 5 中（a）~（d）四段视频均存在不同程度的目标遮挡现象。在正常情况下，三种算法均能对原有目标进行有效跟踪。当目标受到大面积遮挡时，由于 CSK 只采用了简单的灰度特征进行建模，容易受到相似灰度的遮挡物干扰，

从而导致跟踪精度的下降，表现为跟踪框出现严重偏移，如 girl 视频中的 432 帧；KCF 算法依靠多通道局部特征样本的密集型判别策略对短时目标遮挡具有一定的鲁棒性，但由于该方法对模板进行无差异化更新，容易将遮挡物错误引入目标模板而出现跟踪漂移现象，而本文算法能在目标被大面积遮挡情况下，依靠未被遮挡子块跟踪信息对目标进行准确定位，如 freeman4 和 girl 视频。同样当目标被完全遮挡时，由于三种方法均缺乏对目标运动信息的预测能力，导致跟踪框均锁定在置信图峰值最高的遮挡物上，而 CSK 和 KCF 缺乏遮挡检测机制，跟踪模板容易发生漂移，当目标再次出现时，很难对目标再次捕捉，采用遮挡检测的模板更新机制能够很好地抑制目标漂移，能够对目标实现再次捕获，如 football 和 jogging2 视频所示。

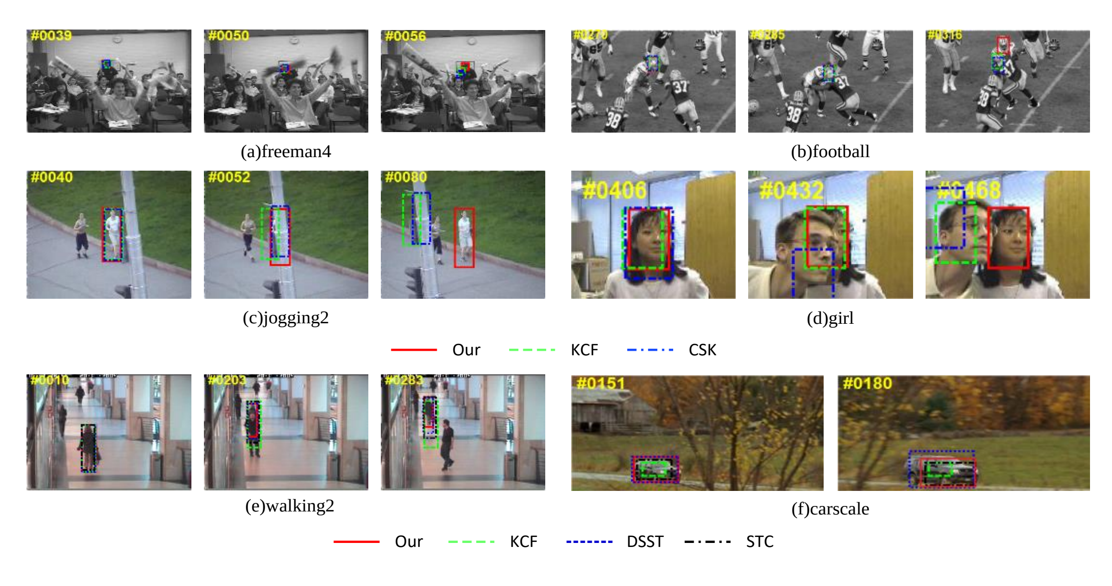


图 5　遮挡及尺度变化场景跟踪结果

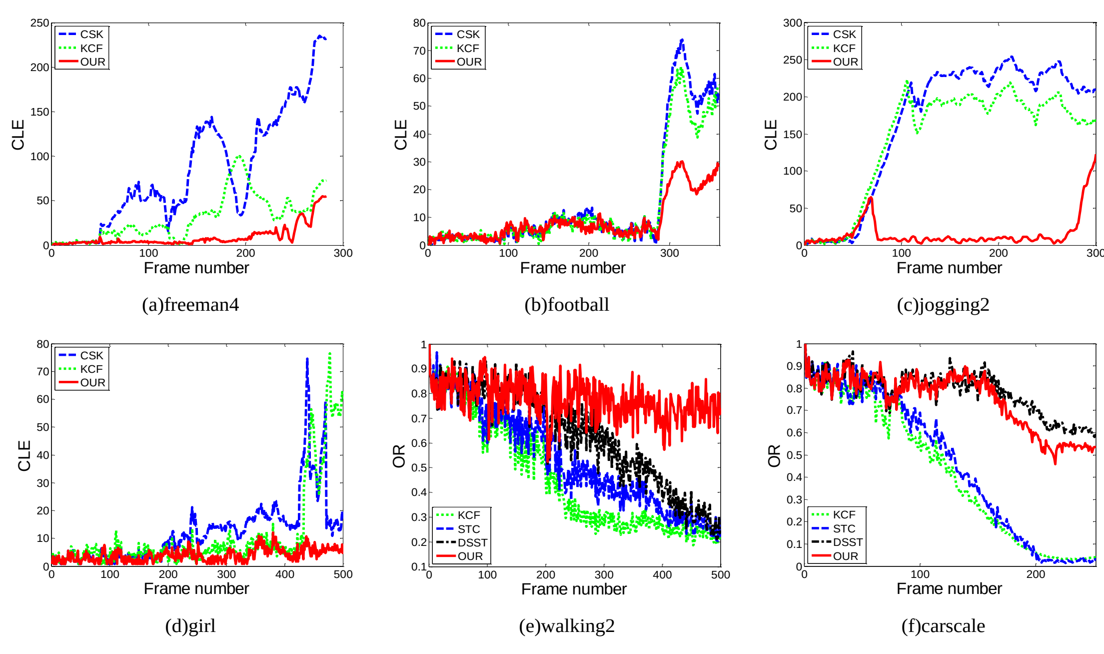


图 6　中心误差及覆盖率分析

尺度变化场景分析：图 5 中的（e）和（f）视频，目标存在明显的尺度变化；在算法选择方面，除了本文方法以及 KCF 外，加入了具备尺度估计能力的相关滤波类算法 STC 及 DSST[16]进行对比分析。从实验结果可以看出，KCF 方法未考虑目标尺度变化，跟踪框大小恒定容易引入过多背景信息或者无法获取目标整体特征而导致跟踪失败，如 walking2 及 carscale 视频序列所示；STC 仅对目标响应的置信图进行估计，当置信图较为可靠时，该方法能够较好地感应目标尺度变化，但因遮挡等原因导致置信图不可靠时，无法对目标的尺度变化作出有效估计。DSST 算法依靠特征金字塔构造目标表观的多尺度样本集，采用尺度相关滤波器实现最佳尺度样本的筛选，但当目标表观出现遮挡时，尺度估计过程容易受到遮挡物干扰，而本文方法则是对多子块置信图以及其各自权重进行联合估计，在一定程度上能减轻遮挡影响，如 walking2 视频中 203 帧到 283 帧所示。

图 6 中（a）~（b）为遮挡类视频中各算法所对应的每帧中心误差曲线。结合图 5 可以看出，当目标出现不同程度遮挡时

（依次对应各视频序列的第 50、270、80 及 432 帧），CSK 及 KCF 跟踪误差出现明显的增大，而本文方法则能保持较好的跟踪精度，但在目标定位策略上采用随机采样的粒子滤波方法进行估计，因此相比其他两种方法跟踪误差上存在较多的小幅度波动现象。图 6（e）（f）为尺度变化场景中各算法所对应的覆盖率曲线。当目标发生尺度变化时，本文尺度估计方法虽比先进的 DSST 略差，但在目标发生部分遮挡情况下（walking2 中第 283 帧），仍能保持较高的覆盖率。

表 1　算法定量对比

算法	Football	Jogging	Girl	Walk2	Carscale	ClifBar	Girl2	Human3	Singer1	Lemming	Freeman4	Faceocc1	平均误差	平均帧率
KCF	40.42	144.99	10.90	12.79	51.72	37.40	87.22	14.36	65.19	83.93	37.72	21.52	50.68	203
CSK	59.08	165.94	19.68	5.53	55.74	68.83	169.77	68.73	82.39	114.68	82.64	24.05	75.59	**286**
DSST	38.73	127.01	11.00	4.17	**21.27**	**34.32**	79.84	19.53	**8.08**	79.84	49.57	21.81	41.26	85
STC	89.97	164.69	23.46	4.63	45.79	84.75	143.21	87.10	35.66	98.30	56.95	86.11	76.72	262
Struck	82.88	120.13	4.24	8.96	80.77	88.11	78.52	56.11	93.20	43.15	41.32	27.88	60.45	21
Our	**22.80**	**29.30**	**4.07**	**3.60**	25.71	123.15	**44.37**	**10.71**	14.50	**32.38**	**28.20**	**19.60**	**29.86**	38

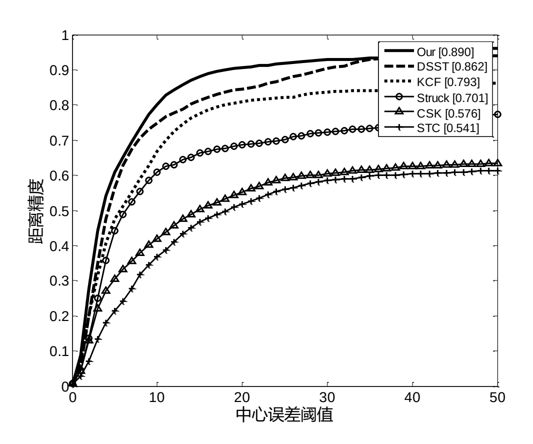
3.3　定量分析实验

表 1 将本文方法与其他几种经典算法（KCF、CSK、STC、DSST、Struck[17]）在 Benchmark 的 12 段含有遮挡或尺度变化类视频序列中，以中心误差为指标进行定量分析。

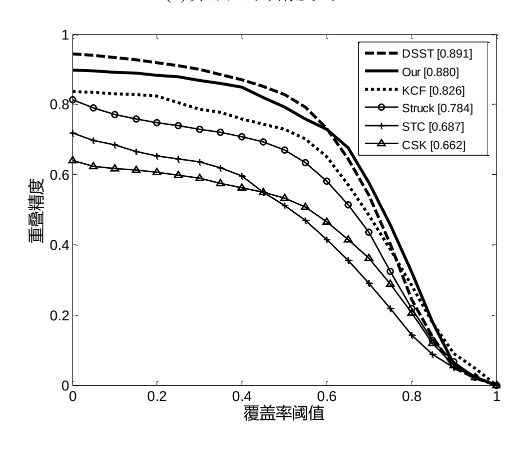
从中可以看出，本文方法在大部分视频序列中跟踪误差较小，并且在 Jogging、Girl2、Walk2 等含有遮挡及尺度放缩场景中仍具有较好的跟踪精度，但本文方法并不能在所有视频中达到最优如 ClifBar，经分析这类视频中目标状态频繁发生较大变化，随着视频帧数的增加，粒子滤波算法在大量迭代过程中可能出现粒子退化而造成无法对目标状态进行有效的估计。算法实时性分析方面，本文方法在分块跟踪的基础上加入粒子滤波置信图估计及遮挡检测模块，虽然提高了遮挡场景中的跟踪精度但也增加了计算复杂度，整体平均速率为每秒 38 帧。图 7（a）（b）分别为各算法的整体距离精度及重叠精度得分曲线。可以看出本文方法与 KCF、DSST 在距离精度上相比分别提高约 10% 和 3%，重叠精度上略低于 DSST。

4　结束语

在核相关滤波跟踪的基础上，本文从多子块联合置信图估计及模板更新两个方面展开相关研究工作。首先依据各子块置信图及其权重生成多子块联合置信图，然后利用粒子滤波算法对联合置信图进行采样，并依据样本框所包围形成的权值密度设定粒子权重，实现目标位置及尺度的最优估计；子块模板更新方面，通过反向投影法将置信度较低的子块映射至目标与背景区域进行遮挡检测，防止因模板错误更新而导致跟踪漂移。在公开测试集中验证并通过定性和定量分析发现，本文方法在目标部分遮挡及尺度放缩场景中仍具有较高的跟踪精度，但在置信图的粒子滤波估计过程中，可能出现因粒子退化而导致的目标状态估计误差增大（图 6（c）），后期可以通过研究更为有效的重采样策略加以改进。



(a)算法距离精度对比



(b)算法重叠精度对比

图 7　算法定量对比分析结果

参考文献：

[1] Ramalakshmi V, Alex M G. Visual object tracking using discriminative correlation filter [C]// Proc of International Conference on Communication and Electronics Systems. 2016: 1-6.

[2] Mehmood R, Nawaz R, Rao N I. Occlusion handling in meanshift tracking using adaptive window normalized cross correlation [C]// Proc of

International Bhurban Conference on Applied Sciences and Technology. 2014: 126-129.

[3] 李伟，曹洁，李军，等. 考虑噪声扰动干扰的室内行人跟踪方法 [J]. 四川大学学报：工程科学版, 2016, 48 (6): 172-179.

[4] Smeulders A W, Chu D M, Cucchiara R, et al. Visual tracking: an experimental survey [J]. IEEE Trans on Pattern Analysis & Machine Intelligence, 2014, 36 (7): 1442-68.

[5] Bolme D S, Beveridge J R, Draper B A, et al. Visual object tracking using adaptive correlation filters [C]// Proc of Computer Vision and Pattern Recognition. 2010: 2544-2550.

[6] Zhang K, Zhang L, Liu Q, et al. Fast visual tracking via dense spatio-temporal context learning [C]// Proc of European Conference on Computer Vision. [S. l.] : Springer International Publishing, 2014: 127-141.

[7] Henriques J F, Rui C, Martins P, et al. Exploiting the circulant structure of tracking-by-detection with kernels [C]// Proc of European Conference on Computer Vision. 2012: 702-715.

[8] Henriques J F, Caseiro R, Martins P, et al. High-speed tracking with kernelized correlation filters [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2015, 37 (3): 583-596.

[9] Yao R, Xia S, Shen F, et al. Exploiting spatial structure from parts for adaptive kernelized correlation filter tracker [J]. IEEE Signal Processing Letters, 2016, 23 (5): 658-662.

[10] 段伟伟，杨学志，方帅，等. 分块核化相关滤波目标跟踪 [J]. 计算机辅助设计与图形学学报, 2016, 28 (7): 1160-1168.

[11] 王暐，王春平，付强，等. 基于分块的尺度自适应 CSK 跟踪算法 [J]. 电光与控制, 2017 (2): 25-29.

[12] Williams C K I. Learning with kernels: support vector machines, regularization, optimization, and beyond [J]. IEEE Trans on Neural Networks, 2005, 16 (3): 781-781.

[13] Lee K, Park S H, Yoon K J. PSR-deterministic search range penalization method on kernelized correlation filter tracker [C]// Proc of International Conference on Ubiquitous Robots and Ambient Intelligence. 2016: 856-860.

[14] 王法胜，鲁明羽，赵清杰，等. 粒子滤波算法 [J]. 计算机学报, 2014, 37 (8): 1679-1694.

[15] Wu Y, Lim J, Yang M H. Object tracking benchmark [J]. IEEE Trans on Pattern Analysis & Machine Intelligence, 2015, 37 (9): 1834.

[16] Danelljan M, Häger G, Khan F S, et al. Accurate scale estimation for robust visual tracking [C]// Proc of British Machine Vision Conference. 2014: 65. 1-65. 11.

[17] Hare S, Saffari A, Torr P H S. Struck: structured output tracking with kernels [C]// Proc of IEEE International Conference on Computer Vision. 2012: 263-270.